Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.6, December 2025, pp. 3509~3517
e-ISSN: 2548-6861

User Requirement Recommendation Model for Waste Reporting Platforms Based on UX Topics and Sentiment Analysis

Irmma Dwijayanti [1]*, Alfirna Rizqi Lahitani [2], Muhammad Habibi [3]*****
* Sistem Informasi, Universitas Jenderal Achmad Yani Yogyakarta
** Teknologi Informasi, Universitas Jenderal Achmad Yani Yogyakarta
***Informatika, Universitas Jenderal Achmad Yani Yogyakarta
irmmadwijayanti@gmail.com [1], alfirnarizqi@gmail.com [2], muhammadhabibi17@gmail.com [3]

Article Info

Article history:

Received 2025-09-27
Revised 2025-11-26
Accepted 2025-11-26

Keyword:

LDA,
Rule-Based Recommendation,
SVM,
UX Honeycomb,
Waste Management.

ABSTRACT

Waste management remains a critical issue in Indonesia, as emphasized in the RPJMN 2025–2029. Ineffective collection and processing services, coupled with limited public participation, continue to hinder progress. Meanwhile, social media has emerged as a primary channel for citizens to express complaints and reports on waste, yet the unstructured nature of comments poses challenges for integration into official reporting systems. This study proposes a user requirements recommendation model based on social media data by integrating sentiment analysis, topic modeling, and rule-based recommendation. Data were collected from YouTube and TikTok comments. Sentiment classification was performed using Support Vector Machine (SVM), while Latent Dirichlet Allocation (LDA) was employed for topic modeling, with results mapped onto the UX Honeycomb dimensions. Recommendation rules were then formulated by combining sentiment polarity with UX dimensions. The SVM model achieved an average accuracy of 87.5% with balanced precision, recall, and F1-score. LDA produced 15 coherent topics, which were distributed across seven UX dimensions. The integration revealed that the main user requirements include transparency in report follow-up through real-time notifications and clear status updates. Additional recommendations involve simplifying the reporting process, providing auto-fill features, improving visual design, and establishing a user appreciation system. The findings demonstrate the potential of leveraging social media comments to systematically capture user requirements, offering practical insights for developers to design waste reporting platforms that are effective, user-friendly, and responsive to community expectations.

I. INTRODUCTION

Waste management is one of the priority focuses in the National Medium-Term Development Plan (RPJMN) 2025–2029, as the implementation of waste management policies in Indonesia remains suboptimal [1]. The lack of infrastructure (such as TPS, TPST, and TPA), low public awareness and participation, as well as inefficient waste collection and processing systems, contribute to waste accumulation, which worsens environmental quality and public health, including the increased risk of flooding caused by clogged urban drainage systems. According to data from the National Disaster Management Agency (BNPB), floods remain the most frequent natural disaster in Indonesia to date [2]. Reinforcing this urgency, a study [3] shows that Indonesian cities continue to struggle with plastic waste. A large portion of waste remains uncollected, and recycling practices still heavily rely on the informal sector. This reality highlights the urgent need for effective waste management services and reporting platforms that can bridge public complaints with government operational responses.

In today's digital era, social media has become a primary channel for the public to voice complaints and reports on waste issues through posts, comments, or replies across platforms such as Instagram, Twitter, TikTok, and YouTube [4]. Previous studies have indicated that these channels are effective in capturing public perceptions and attention toward environmental issues [5]. However, such reports are unstructured and not integrated into official reporting systems, making it difficult to compile and follow up consistently and promptly across authorities.

Several studies have shown that social media data can be utilized to understand public perception, identify key issues, and reveal important themes. Study [6] classified public comments on YouTube regarding waste issues in Yogyakarta to examine community perceptions, yielding a significant distribution of positive, neutral, and negative sentiments. Another study [7] employed the Support Vector Machine (SVM) algorithm to analyze public opinion on Twitter regarding waste, showing that negative sentiments outweighed positive ones, reflecting public dissatisfaction. Nevertheless, the use of social media to build an integrated waste reporting platform that directly channels community reports to the authorities remains limited.

Designing a user-centered platform has the potential to significantly increase adoption and public participation. According to ISO 9241-210, User-Centered Design (UCD) is a design approach that focuses on users by placing their needs, characteristics, and limitations at the center of the entire system development process [8]. UCD is closely tied to the concept of User Experience (UX), which refers to the overall quality of experience perceived by users when interacting with a product or service [9]. Therefore, when building a platform aligned with user requirements, the first essential step is to capture those needs. Large-scale data directly sourced from public interactions on social media can be leveraged to identify real community needs, ranging from ease of use and response speed to transparency in report follow-ups.

Topic modeling is one of the techniques that can be used to analyze unstructured data such as social media comments. Previous studies indicate that Latent Dirichlet Allocation (LDA) remains a robust method for extracting themes or topics from reviews or comments, particularly for medium-sized unstructured corpora [10]. Compared to other methods such as NMF [11], LSA [12], BERTopic [13], or ensemble models [14], methodological comparisons show that each technique has its own advantages and limitations, yet LDA remains competitive, especially when topic coherence is quantitatively evaluated using coherence metrics.

In addition, sentiment analysis is a widely used technique for analyzing comments. SVM remains relevant for short texts such as social media content or reviews and often outperforms classical methods like Logistic Regression [15], and Naive Bayes [16] in terms of accuracy and F1-score [17]. Comparative and applied studies on reviews from applications and social media platforms demonstrate that SVM is stable, replicable, and a reliable choice when feature clarity and development speed are primary considerations [18].

However, the outputs of topic modeling and sentiment analysis are still mostly used as descriptive analytics rather than as a formal basis for operational recommendation rules that can be directly implemented. No research has been found that explicitly integrates the output of UX aspect-based topic models and sentiment classification to build operational rule-based recommendation models in the context of waste reporting platforms or public services. Some studies have combined sentiment with non-textual signals for recommendation purposes [19] or integrated topics and sentiments to generate strategic insights [20], but not to establish operational policy or feature rules that can be directly applied in public reporting platforms.

Therefore, this study aims to build a user needs recommendation model based on social media data by integrating sentiment analysis, UX-based topic modeling, and rule-based recommendations to generate a list of platform needs. The main contribution of this study lies in the development of the model as a conceptual methodology that is not yet implemented, but has strong potential for application in the development of real systems. This model is intended as a basis for designing new applications that utilize social media data sources to build a waste reporting platform that is relevant to user needs, while also supporting the government, especially local governments, in handling and following up on reports of waste problems in their regions more effectively.

II. METHOD

This research consists of data collection, data preprocessing, data labeling, SVM sentiment analysis, UX topic modelling using LDA, and rule-based implementation. The novelty of this research lies in the integration of these three methods, where topic modeling results are interpreted based on UX aspects to produce a recommendation model that is not only data-driven but also oriented towards the user's real experiences and needs. The research phases are illustrated in Figure 1.

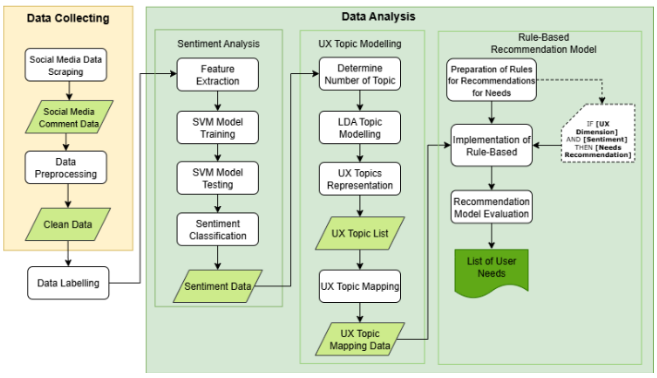

Figure 1. Research Phase

A. Data Collection

Data were collected by scraping comments from YouTube, YouTube Shorts, and TikTok using keywords related to "pelaporan sampah" and "pengelolaan sampah". The selection of video-based platforms was considered appropriate because comments on video content tend to be more specific and contextual to the topics presented, thus providing more relevant data compared to general text-based social media. From this scraping process, 19015 comments were obtained as a research dataset and all data used is public and has been anonymized to ensure compliance with user ethics and privacy.

B. Data Preprocessing

The preprocessing stage was carried out to ensure that the data were clean and ready for analysis. Several steps were applied, including case folding, which converts all text into lowercase for consistency; punctuation and number removal, which eliminates punctuation marks, numbers, and symbols that do not contribute meaningful information; and stopword removal, which removes common words such as "yang," "dan," or "di" that are irrelevant for analysis. The preprocessing stage produced a clean dataset that served as input for labelling and subsequent analysis.

C. Data Labelling

Data labeling was conducted to prepare training and testing datasets for sentiment analysis. The method applied was lexicon-based labelling using InSet (Indonesian Sentiment Lexicon), which contains 3609 positive words and 6609 negative words with weighted scores ranging from -5 to +5 [21]. The labelling process was performed automatically by assigning a positive or negative label to each comment based on the cumulative score of words contained in the InSet. As a result, 9772 comments were labelled as positive and 9243 comments as negative.

D. Data Analysis

1) *Sentiment Analysis:* Sentiment analysis was conducted using the Support Vector Machine (SVM) algorithm. SVM is a supervised learning classification algorithm that separates two classes through an optimal hyperplane, with the margin defined as the shortest distance between the nearest data points to the hyperplane [22]. In this study, user comments were classified into two sentiment classes: positive and negative. Mathematically, SVM seeks to maximize the margin between two classes with the following function 1:

$$f(x) = w.x + b \quad (1)$$

where w represents the weight vector, x the input vector, and b the bias [23]. This model was selected due to its proven effectiveness in natural language sentiment analysis tasks.

2) *UX Topic Modeling:* The next stage involved topic modeling using LDA, an unsupervised learning method that represents documents as a mixture of topic distributions [24]. The optimal number of topics was determined using coherence scores to ensure that the topics were semantically coherent. Typically, LDA outputs do not have explicit labels, requiring human interpretation to assign meaning to each topic group based on its constituent words.

This study interprets the meaning of each topic group based on its constituent words, then maps each topic into UX Honeycomb dimensions to become a specific UX topic. The mapping process is carried out by analyzing a list of topics generated from the best number of topics, where each topic has a set of constituent words used to interpret its meaning. Usually, the determination of topic meaning is done manually by humans, but in this study the process is specified to the UX context. Each topic is mapped to a UX Honeycomb aspect based on the match between the constituent words of the topic and a list of keywords that represent each UX aspect — namely usable (ease of use), useful (usefulness), desirable (aesthetics), findable (ease of information search), accessible (accessibility), credible (trustworthiness), and valuable (value) [25]. Thus, each topic in the modeling results list will be classified into one of the UX dimensions, and this mapped topic list is then used to determine the UX aspect of each comment in the entire data. Each topic was associated with 50 keywords related to these dimensions in the context of a waste management platform, compiled through a literature review. Table I presents the list of 20 keywords for each UX topic.

TABLE I
UX TOPICS KEYWORD LIST

UX Topic	Keyword	Ref
Usable	mudah, susah, ribet, gampang, jelas, cepat, lemot, error, lag, loading, lambat, lancar, stabil, berat, ringan, simpel, bingung, tombol, klik, menu	[26]
Useful	lapor, sampah, foto, lokasi, alamat, titik, buang, penuh, kotor, bersih, angkut, petugas, jadwal, update, status, progres, konfirmasi, keluhan, aduan, masalah	[27]
Desirable	bagus, jelek, keren, modern, jadul, rapi, acak, warna, estetik, tampilan, ikon, font, huruf, gelap, terang, tema, style, desain, animasi, visual	[28]
Accesible	akses, sinyal, offline, online, kuota, hp, android, ios, layar, kecil, besar, teks, huruf, zoom, warna, kontras, terang, gelap, mode, malam	[27]
Findable	cari, nemu, ketemu, lokasi, alamat, titik, map, peta, gps, arah, filter, kategori, pilih, urut, sort, tag, label, info, daftar, riwayat	[29]
Credible	resmi, asli, palsu, bohong, benar, salah, aman, valid, privasi, bocor, percaya, jujur, jelas, transparan, update, telat, waktu, bukti, instansi, pemerintah	[27]
Valuable	manfaat, berguna, nolong, sia-sia, hasil, beres, dampak, puas, kecewa, senang, lega, perubahan, solusi, kontribusi, untung, rugi, progres, tindak, respon, efisien	[27]

3) *Implementasi Rule-Based:* The final stage was developing user requirements recommendations through a Rule-Based Recommendation (RBR) approach. RBR applies a set of explicit IF–THEN rules to generate recommendations [30]. Each rule clearly specifies conditions (IF) and corresponding actions (THEN), making the system both interpretable and explainable.

```
Rule 1 : Rule-based UX Recommendation
if UX_Dimension = Usable and Sentiment = Negative then
    Recommendation ← "Simplify and make the reporting flow more intuitive
    (e.g., using buttons/icons instead of long text)"
end if
```

Figure 2. Rule-based UX Recommendation

In this study, each rule connects the sentiment type (positive/negative), the related UX topic, and the corresponding user need. Based on these three aspects, rules are formulated to produce a structured list of user needs, as shown in Figure 2. Table II lists the rules obtained from the literature review, which illustrate the relationship between UX topics, sentiment polarity, and user need recommendations generated for the waste reporting platform. User needs recommendations are generated from the integration process of sentiment analysis results using SVM and topic modeling using LDA, which is done by combining sentiment-classified comment data with the results of UX topic mapping and a compiled list of rules. Each comment is then matched between its sentiment value and UX topic to generate user need recommendations based on the list of rules.

TABLE II
RULES LIST

ID	UX Topic	Sentiment	Recommendation	Ref
R1	Usable	Negative	Simplify and make the reporting flow more intuitive (e.g., using buttons/icons instead of long text)	[31]
R2	Usable	Positive	Add an auto-save feature to prevent data loss when the application is closed.	[31]
R3	Useful	Negative	Add a report status feature (e.g., "in progress," "completed")	[31]
R4	Useful	Positive	Add real-time notifications for report follow-up updates	[31]
R5	Desirable	Negative	Improve visual design with consistent icons, eye-friendly colors, and a modern layout	[8], [31]
R6	Desirable	Positive	Provide light/dark theme options for user visual preferences	[8], [31]
R7	Findable	Negative	Make report categories clearer and add a report search feature based on keywords or locations	[29]
R8	Findable	Positive	Add a report filter feature (e.g., by region or waste type)	[29]
R9	Accessibl e	Negative	Add alt-text, screen reader support, and proper color contrast	[32]
R10	Accessibl e	Positive	Add voice input support to simplify reporting	[32]
R11	Credible	Negative	Display the name of the institution or official handling the report	[33]
R12	Credible	Positive	Add verification badges for reports already processed by official institutions	[34]
R13	Valuable	Negative	Introduce a reward system (points/badges) to make users feel their contributions are valuable	[27]
R14	Valuable	Positive	Display user contribution statistics (e.g., amount of waste handled through their reports)	[27]

III. RESULTS & DISCUSSIONS

A. *Sentiment Analysis*

The preprocessed comment dataset was used for sentiment classification using the SVM method. In this sentiment analysis, the dataset was split into 70% training data and 30% testing data, consisting of 13310 comments for training, and 5705 comments for testing. To avoid overfitting and ensure performance consistency, 10-fold cross-validation was applied to the training data.

The results of each fold are presented in Figure 3. From the 10 iterations, the model achieved an average accuracy of 0.875, precision of 0.877, recall of 0.875, and F1-score of 0.875. These values indicate that the SVM model is relatively stable with consistent performance across folds. The best-performing model from training was then evaluated using the testing data. The evaluation results based on accuracy, precision, recall, and F1-score are shown in Table III.

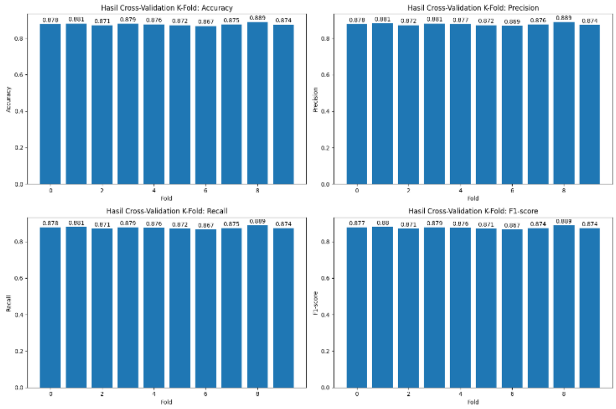

Figure 3 10-fold Cross Validation Results

TABLE III
RESULT OF MODEL EVALUATION

Metric	*Accuracy*	*Precision*	*Recall*	*F-1 Score*
Value	0.883	0.884	0.883	0.883

The model demonstrated promising performance with consistent scores across all metrics. The similarity of results between cross-validation and testing suggests that the model is capable of effective classification.

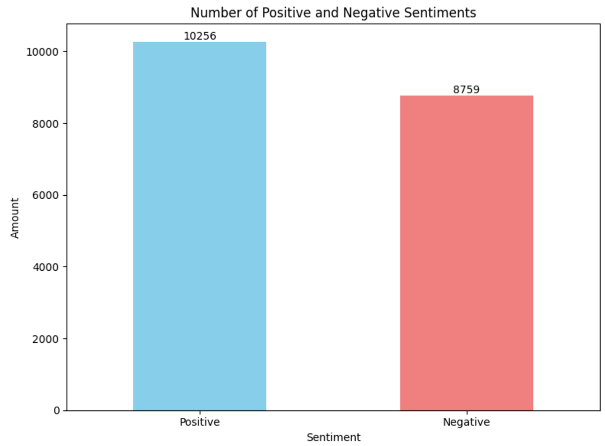

Figure 4 Distribution of Positive and Negative Sentiments

Subsequently, the trained SVM model was used to classify the entire dataset of comments into positive and negative labels. The sentiment distribution is illustrated in Figure 4. The classification results show that positive sentiment is more dominant than negative sentiment. Thus, it can be concluded that public sentiment regarding waste issues tends to be more positive overall, although negative sentiment still appears frequently in discussions.

B. *UX Topic Modeling*

The LDA topic modeling process began by calculating coherence scores to determine the optimal number of topics.

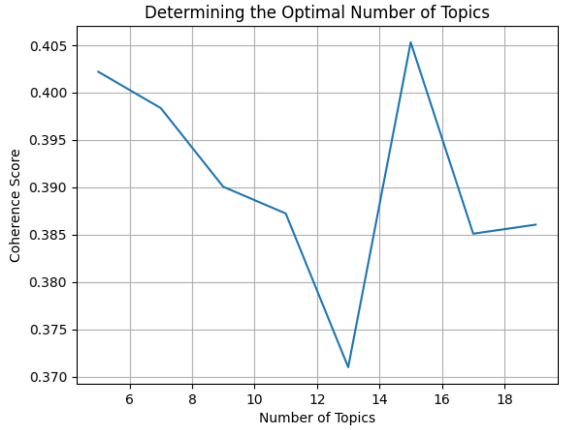

Figure 5 Coherence Score

As shown in Figure 5, the optimal number of topics was found to be 15, with a coherence score of 0.4053. Each topic consisted of keywords that could be used for topic interpretation. Table IV presents the list of 15 topics along with the top 10 contributing keywords for each.

TABLE IV
TOPIC LIST

Topic no.	**Words of Topic**	**UX Mapping**
1	kebersihan, sampahnya, petugas, mentri, iman, duitnya, smpah, lahan, dibakar, berharap	Useful
2	gubernur, membuang, dibersihkan, dibuang, terbaik, giliran, respect, hulu, jokowi, sungainya	Credible
3	daur, organik, menjaga, membersihkan, mudah, memilah, uangnya, apapun, pupuk, kasian	Usable
4	kdm, kerja, jabar, disana, ton, korup, tps, dek, diolah, jepang	Uncategorized
5	kesadaran, pemimpin, dedi, warganya, numpuk, jorok, harga, kab, mengurangi, disiplin	Desirable
6	semoga, pandawara, bersihkan, kesehatan, dijadikan, meniru, kebaikan, mengelola, aliran	Useful
7	pemerintah, gebang, bantar, pemerintahan, swasta, menggunung, nasional, masyaallah, tpst, tangga	Useful
8	sehat, dedy, videonya, sayang, lanjutkan, omon, peraturan, pengusaha, urusan, dicontoh	Uncategorized
9	bupati, anggaran, pembuangan, konoha, kebanyakan, gub, proyek, jin, perubahan, walikota	Valuable
10	gini, alhamdulillah, rk, pembakaran, hasilnya, industri, menumpuk, ekspor, pemerintahnya, membangun	Valuable
11	banjir, bekasi, owen, bali, citarum, cuan, kamil, tempatnya, pemda, ridwan	Uncategorized
12	pengolahan, pejabat, belajar, generasi, limbah, emas, terapkan, kecamatan, pemulung, berguna	Valuable

Topic no.	Words of Topic	UX Mapping
13	masyarakatnya, makan, mengolah, jadikan, ahok, aturan, koruptor, singapura, pencitraan, kebijakan	Usable
14	lingkungan, pengelolaan, tpa, bermanfaat, wajib, namanya, kerjanya, pengen, positif, mencontoh	Uncategorized
15	sembarangan, kang, aing, setuju, membantu, import, periode, perduli, disitu, husein	Useful

The list of topics was then mapped into the UX Honeycomb dimensions by matching the dominant keywords in each topic with indicators from the seven UX dimensions, namely Usable, Useful, Desirable, Accessible, Findable, Credible, and Valuable. For example, Topic 3 has the topic-forming words "mudah, membersihkan, memilah" which match the keywords in the *Usable* dimension. It can be seen that the results of the UX topic mapping do not appear in all dimensions; if there is no match, it is categorized as Uncategorized.

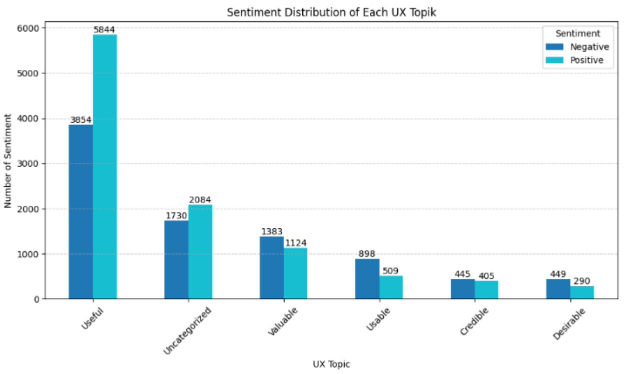

Figure 6 Sentiment Distribution per UX Topic

The overall comment data is automatically classified into appropriate topics based on a list of previously mapped topics to UX dimensions. This process is performed by matching the words in each comment with the constituent words in each LDA topic that have been labeled with UX dimensions, so that each comment can be identified as the most relevant UX topic.

Figure 6 illustrates the distribution of comments across UX topics along with their sentiment polarity. It can be observed that the **Useful** topic was the most frequently discussed aspect, with positive sentiment dominating over negative sentiment. This indicates that the public places significant emphasis on the functionality and usefulness of the waste reporting platform, expressed both through positive appreciation and complaints regarding service effectiveness. This topic thus emerges as the top priority, given its substantially higher volume compared to other topics. The **Valuable** topic ranked third, with negative sentiment outweighing positive sentiment. This suggests that while users recognized the tangible benefits of the platform, they also voiced considerable criticism regarding its added value. In the **Usable** dimension, negative sentiment was also more dominant than positive sentiment, highlighting that ease of use remains a critical challenge, with complaints outnumbering appreciation. Meanwhile, the **Credible** and **Desirable** dimensions were less frequently discussed but remain important, as they relate to trust in the platform and its visual appeal.

Overall, these findings show that the public discussion is most heavily centered on the **Useful** dimension as the core of user experience, while at the same time, the **Usable** dimension also requires significant attention due to its higher share of negative feedback. This distribution underscores that although positive sentiment is relatively high, issues related to usability and service effectiveness remain major concerns in platform development.

In addition to the six UX topics, the analysis also identified a group of comments categorized as **Uncategorized**. This category emerged because certain comments did not contain relevant keywords that could be mapped to the defined UX dimensions. In other words, the themes expressed in these comments were not directly related to usefulness, usability, desirability, value, credibility, or accessibility, and thus could not be classified into the existing UX dimensions.

Figure 7 WordCloud Uncategorized Categories

The word cloud visualization in Figure 7 highlights dominant keywords such as *governor, related, indifferent, opportunity, habit, energy, investment, YouTube,* and *sea.* These words indicate that some comments discussed issues outside the core focus of user experience, including politics, local government policies, societal habits, investment opportunities, and broader discussions on the environment and energy. This confirms that while the data collection targeted waste-related issues, public discussions often expanded into broader social, economic, and political topics.

C. Rule-Based Recommendation Model

The results of the comment analysis, which had been mapped into specific UX topics and classified by positive and negative sentiment, were used to develop the RBR model. Each comment was analyzed using the predefined rules, where each rule represents a combination of UX topics and sentiment polarity.

Through the application of RBR, every comment was not only mapped to its UX topic and sentiment but also automatically generated a corresponding recommendation. All recommendations were then aggregated to observe the distribution of comments supporting each identified need. Consequently, the outcome was a prioritized list of user requirements, measured by the number of comments underlying each need, as presented in Table V.

TABLE V
USER RECOMMENDATIONS LIST

Rule ID	Recommendation	Amount
R4	Add real-time notifications for report follow-up updates	5844
R3	Add a report status feature (e.g., "in progress," "completed")	3854
R13	Introduce a reward system (points/badges) to make users feel their contributions are valuable	1383
R14	Display user contribution statistics (e.g., amount of waste handled through their reports)	1383
R1	Simplify and make the reporting flow more intuitive (e.g., using buttons/icons instead of long text)	898
R2	Add an auto-save feature to prevent data loss when the application is closed.	509
R5	Improve visual design with consistent icons, eye-friendly colors, and a modern layout	449
R11	Display the name of the institution or official handling the report	445
R12	Add verification badges for reports already processed by official institutions	405
R6	Provide light/dark theme options for user visual preferences	209

Based on the analysis results, the requirement with the highest number of comments was the addition of real-time notifications to update report follow-ups (R4). This highlights that users place strong emphasis on transparency and timely feedback from the reporting management system. The next most significant requirement was the addition of a report status feature (R3), underscoring the importance of providing clear information about report progress. This need was further reinforced by user requests to display details of the authorities handling the reports (R11, R12). Reward-related features that acknowledge user contributions (R13, R14) also attracted notable attention, indicating that users value systems that appreciate and recognize their participation. Other requirements, though supported by fewer comments, were still considered relevant. These include simplifying the reporting flow (R1), adding an auto-fill feature for report data (R2), and improving the visual design to be more modern and user-friendly (R5, R6).

However, not all of the designed rules (R1–R14) appeared in the final distribution of comments. In particular, rules related to the **Findable** dimension (R7–R8) and the **Accessible** dimension (R9–R10) were not significantly identified in the comment data. This suggests that aspects such as information searchability and accessibility have not yet emerged as major public concerns in the context of waste reporting.

D. Discussion

This study integrates sentiment analysis using SVM with topic modeling via LDA, which were then combined into a RBR model. From the algorithmic perspective, SVM proved to be reliable in classifying public sentiment within large and diverse comment datasets, as demonstrated by an average accuracy of 0.875 in 10-fold cross-validation. The relatively balanced precision, recall, and F1-scores further confirm the model's consistency in distinguishing between positive and negative comments.

Subsequently, the LDA algorithm was applied to identify topic structures within the comment data. Coherence score calculations indicated that the optimal number of topics was 15, with a score of 0.4053. These topics were mapped onto the seven UX Honeycomb dimensions, enabling each comment to be linked not only to sentiment polarity but also to a relevant UX topic. The use of UX Honeycomb in evaluating digital public services has also been highlighted in prior studies, showing its effectiveness in identifying user priorities across dimensions such as usefulness and trust [35]. This mapping facilitated deeper insights into the aspects of user experience most frequently discussed by the public.

The integration of both models resulted in the RBR framework, where each combination of UX topic and sentiment was translated into a specific user requirement. This emphasizes that users highly value quick responses and clarity in the reporting process. The findings show that the most dominant needs were transparency in report follow-ups through real-time notifications and the availability of a report status feature. This resonates with earlier e-government research emphasizing transparency and responsiveness as critical factors for building user trust [36]. Additional requirements, such as displaying responsible authorities and incorporating a reward system for contributors, also emerged as important, reflecting the role of credibility and recognition in fostering public trust. Similar approaches have been examined in waste management studies, where incentive and reward mechanisms were found to significantly increase community participation in recycling programs [37].

Although less frequent, other UX dimensions, such as Usable, Desirable, Findable, and Accessible, also provided valuable insights. Suggestions related to simplifying the reporting flow, adding auto-fill features, and improving visual design highlighted the importance of ease of use and aesthetic appeal. Previous UX evaluations of civic applications also pointed out that usability and visual design strongly influence adoption and sustained use [35]. Meanwhile, rules associated with information findability and accessibility did not appear significantly in the comment data, suggesting that users currently prioritize transparency and responsiveness over these aspects. Nonetheless, the Findable and Accessible dimensions should remain a focus in future development to ensure platform inclusivity for diverse user requirements.

The use of RBR in this study offers several advantages. First, the rules are transparent and interpretable, allowing researchers and decision-makers to easily trace the rationale behind each recommendation. Second, the method is flexible for incorporating domain knowledge, as rules can be tailored to UX Honeycomb indicators and sentiment analysis outputs. Third, RBR effectively translates unstructured data (comments) into concrete user requirement recommendations.

However, this approach also has limitations. One limitation is its reliance on the completeness and quality of the rules: if the rules fail to cover certain keyword variations or contexts, some user requirements may go undetected. Furthermore, rule-based systems tend to be less adaptive to new data patterns compared to purely machine learning–based approaches. For this reason, future research could extend the model by combining rule-based methods with automated learning techniques to enhance flexibility and coverage.

Overall, the integration of SVM, LDA, and RBR in this study demonstrates that social media comments can be leveraged to extract real community needs. The developed model is able to process unstructured comment data into measurable information regarding UX dimensions and relevant sentiments. The results provide a systematic overview of user needs, such as transparency in report follow-up, providing clear report status, and strengthening user credibility and appreciation. Compared to non-UX-based needs gathering approaches or manual needs exploration, which tend to be subjective, time-consuming, and rely on limited interviews or observations, this automated UX-based approach offers a much broader scope and is based on real data on user behavior.

These findings confirm that social media comment analysis not only reflects public perception but can also serve as the basis for recommendations for developing a user-centered waste reporting platform. At the same time, usability and accessibility aspects need to be a long-term development agenda to ensure the platform is truly inclusive, effective, and trusted by the public. Furthermore, while social media comment analysis provides a broad picture of user needs, this data source is subject to potential biases, such as spam comments, the use of sarcasm, which is difficult to detect with sentiment analysis, and the dominance of negative comments, which can influence the distribution of recommendations. While some irrelevant comments are successfully filtered through the Uncategorized category, the natural bias of social media remains a limitation that must be considered when interpreting the results.

This model can be practically implemented by embedding it directly into a waste reporting platform. The system can automatically process user comments or feedback and then generate recommendations for needs or feature improvements in real time. This allows the platform to adapt more quickly to community needs and support responsive and sustainable service development.

IV. CONCLUSION

This study developed a user requirement recommendation model based on social media data by integrating sentiment analysis using SVM, topic modeling with LDA mapped to UX dimensions, and a rule-based recommendation approach. The model systematically generated user requirements from public comments, with the top priorities identified as transparency in report follow-ups, clarity of report status, and enhanced credibility and user recognition. The proposed model provides a structured reference for developers in designing waste reporting platforms that are functional, user-friendly, and aligned with community needs. For future work, the model can be extended by incorporating deep learning approaches or hybrid recommendation methods to improve adaptability, as well as testing the framework on broader datasets and other domains beyond waste reporting.

ACKNOWLEDGMENT

This acknowledgment is addressed to the Directorate of Research and Community Service (DP2M), Ministry of Higher Education, Science, and Technology (Kemendiktisaintek) of the Republic of Indonesia, for providing the opportunity to conduct this research in 2025 under the New Lecturer Research (PDP) scheme (Contract Number: SPK/132/LPPMUNJAYA/VI/2025).

REFERENCES

[1] Jaringan Dokumentasi dan Informasi Hukum Kementerian Sekretaris Negara, *Peraturan Presiden (Perpres) tentang Rencana Pembangunan Jangka Menengah Nasional Tahun 2025 - 2029*. Indonesia: Jaringan Dokumentasi dan Informasi Hukum Kementerian Sekretaris Negara, 2025. [Online]. Available: https://jdih.setneg.go.id/

[2] B. R. Alfathi, "Banjir Dominasi Bencana Alam Indonesia 2024," https://data.goodstats.id/. Accessed: Apr. 02, 2025. [Online]. Available: https://data.goodstats.id/statistic/banjir-dominasi-bencana-alam-indonesia-2024-DH6lL

[3] Y. Zahrah, J. Yu, and X. Liu, "How Indonesia's Cities Are Grappling with Plastic Waste: An Integrated Approach towards Sustainable Plastic Waste Management," *Sustain.* , vol. 16, no. 10, p. 3921, May 2024, doi: 10.3390/SU16103921/S1.

[4] J. S. Maheswari, E. S. Hikmah, and M. B. Rizaldhi, "Penggunaan Media Instagram dalam Kampanye Pengurangan Sampah Plastik: Studi Pustaka Artikel Ilmiah Periode 2019-2022," in *Prosiding Seminas Nasional Universitas Negeri Surabaya*, Surabaya: Universitas Negeri Surabaya, 2023, pp. 702–711.

[5] E. S. Aulia, N. A. Putri, A. T. Lumintang, N. Nabiilah, B. S. Inessafitri, and D. E. Rahmawati, "Sentiment Analysis: Proses Komunikasi Pemerintahan Pada Gerakan Zero Sampah Anorganik Di Kota Yogyakarta Menuju Kota Berkelanjutan," *JIIP J. Ilm. Ilmu Pemerintah.*, vol. 9, no. 2, pp. 169–180, Oct. 2024, doi: 10.14710/JIIP.V9I2.24032.

[6] H. Yuliansyah, S. A. Mulasari, S. Sulistyawati, F. A. Ghozali, and B. Sudarsono, "Sentiment Analysis of the Waste Problem based on YouTube comments using VADER and Deep Translator," *J. MEDIA Inform. BUDIDARMA*, vol. 8, no. 1, p. 663, Feb. 2024, doi: 10.30865/MIB.V8I1.6918.

[7] R. A. Nugroho, S. H. Wijono, K. Pinaryanto, R. Gunawan, and F. X. Sinungharjo, "Sentiment Analysis on Tweets about Waste Problem in Yogyakarta using SVM," *Int. J. Appl. Sci. Smart Technol.*, vol. 6, no. 1, pp. 183–196, Jun. 2024, doi: 10.24071/IJASST.V6I1.7415.

[8] "ISO 9241-210:2019 - Ergonomics of human-system interaction —

Part 210: Human-centred design for interactive systems." Accessed: Sep. 18, 2025. [Online]. Available: https://www.iso.org/standard/77520.html

[9] Interaction Design Foundation-IxDF, "User Experience Design," Interaction Design Foundation. Accessed: Sep. 19, 2025. [Online]. Available: https://www.interaction-design.org/literature/topics/ux-design

[10] N. N. Hidayati and A. Parlina, "Performance Comparison of Topic Modeling Algorithms on Indonesian Short Texts," *ACM Int. Conf. Proceeding Ser.*, pp. 117–120, Nov. 2022, doi: 10.1145/3575882.3575905.

[11] S. Mifrah and E. H. Benlahmar, "Topic Modeling Coherence: A Comparative Study beetween LDA and NMF Models using COVID'19 Corpus," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 5756–5761, 2022, Accessed: Sep. 19, 2025. [Online]. Available: https://www.warse.org/IJATCSE/static/pdf/file/ijatcse231942020.pdf?

[12] A. Rkia, A. Fatima-Azzahrae, A. Mehdi, and L. Lily, "NLP and Topic Modeling with LDA, LSA, and NMF for Monitoring Psychosocial Well-being in Monthly Surveys," *Procedia Comput. Sci.*, vol. 251, pp. 398–405, Jan. 2024, doi: 10.1016/J.PROCS.2024.11.126.

[13] E. Chagnon, R. Pandolfi, J. Donatelli, and D. Ushizima, "Benchmarking topic models on scientific articles using BERTeley," *Nat. Lang. Process. J.*, vol. 6, p. 100044, Mar. 2024, doi: 10.1016/J.NLP.2023.100044.

[14] K. Kusumaningtyas, A. R. Lahitani, I. Dwijayanti, M. Habibi, and A. Yani Yogyakarta, "Mapping User Dissatisfaction in Mobile Banking Applications Using Ensemble Clustering LDA and LSA," *IJNMT (International J. New Media Technol.*, vol. 12, no. 1, pp. 32–38, Jun. 2025, doi: 10.31937/IJNMT.V12I1.3804.

[15] D. Y. Saraswati, M. R. Handayani, K. Umam, and M. I. Mustofa, "Sentiment Classification of MyPertamina Reviews Using Naïve Bayes and Logistic Regression," *J. Appl. Informatics Comput.*, vol. 9, no. 4, pp. 1319–1325, Aug. 2025, doi: 10.30871/JAIC.V9I4.9723.

[16] S. Prayudani, D. R. B. Situmorang, R. Hidayah, and H. S. Ginting, "Sentiment Analysis of Social Media X in the 2024 Indonesian Presidential Election Using the Naive Bayes Algorithm: Candidates' Backgrounds and Political Promises," *J. Appl. Informatics Comput.*, vol. 8, no. 2, pp. 291–295, Oct. 2024, doi: 10.30871/JAIC.V8I2.7580.

[17] J. O. Leandro and M. I. Fianty, "Evaluation of Sentiment Analysis Methods for Social Media Applications: A Comparison of Support Vector Machines and Naïve Bayes," *JOIV Int. J. Informatics Vis.*, vol. 9, no. 2, pp. 796–807, Mar. 2025, doi: 10.62527/JOIV.9.2.2905.

[18] Natasha and R. R. Suryono, "Sentiment Analysis of the Influence of the Korean Wave in Indonesia using the Naive Bayes Method and Support Vector Machine," *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 1, pp. 308–319, Mar. 2025, doi: 10.35314/85X4WD90.

[19] M. Elahi, D. Khosh Kholgh, M. S. Kiarostami, M. Oussalah, and S. Saghari, "Hybrid recommendation by incorporating the sentiment of product reviews," *Inf. Sci. (Ny).*, vol. 625, pp. 738–756, May 2023, doi: 10.1016/J.INS.2023.01.051.

[20] Y. P. Mulyani *et al.*, "Analyzing public discourse on photovoltaic (PV) adoption in Indonesia: A topic-based sentiment analysis of news articles and social media," *J. Clean. Prod.*, vol. 434, p. 140233, Jan. 2024, doi: 10.1016/J.JCLEPRO.2023.140233.

[21] F. Koto and G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," *Proc. 2017 Int. Conf. Asian Lang. Process. IALP 2017*, vol. 2018-January, pp. 391–394, Jul. 2017, doi: 10.1109/IALP.2017.8300625.

[22] K. Kusumaningtyas, M. Habibi, I. Dwijayanti, and R. Sumiyarini, "Tweet Analysis of Mental Illness Using K-Means Clustering and Support Vector Machine," *Telematika*, vol. 20, no. 3, p. 295, 2023, doi: 10.31315/telematika.v20i3.9820.

[23] W. Widayani and H. Harliana, "Analisis Support Vector Machine Untuk Pemberian Rekomendasi Penundaan Biaya Kuliah Mahasiswa," *J. Sains dan Inform.*, vol. 7, no. 1, pp. 20–27, Jun. 2021, doi: 10.34128/JSI.V7I1.268.

[24] D. M. Blei, A. Y. Ng, and J. B. Edu, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003, doi: 10.5555/944919.944937.

[25] P. Morville, "User Experience Design," Semantic Studios. Accessed: Sep. 19, 2025. [Online]. Available: http://semanticstudios.com/user_experience_design/

[26] S. P. Handasari, R. Wulandari, and Haikal, "Evaluation of the Usability and User Experience of the Jaminan Kesehatan Nasional Mobile Application in Indonesia," *Healthc. Inform. Res.*, vol. 30, no. 4, p. 324, Oct. 2024, doi: 10.4258/HIR.2024.30.4.324.

[27] R. Bastardo, J. Pavão, and N. P. Rocha, "Methodological Quality of User-Centered Usability Evaluation of Digital Applications to Promote Citizens' Engagement and Participation in Public Governance: A Systematic Literature Review," *Digit. 2024, Vol. 4, Pages 740-761*, vol. 4, no. 3, pp. 740–761, Sep. 2024, doi: 10.3390/DIGITAL4030038.

[28] W. W. Sejati, E. Sibagariang, and H. Setiawan, "EVALUATION AND DESIGN OF USER EXPERIENCE IN COVE APPLICATION USING UX HONEYCOMB," *Comput. Based Inf. Syst. J.*, vol. 12, no. 2, pp. 29–39, Sep. 2024, doi: 10.33884/CBIS.V12I2.9200.

[29] Y. Prasetya and R. Rahmi, "Pilot Study on User Experience Analysis of Universitas Indonesia Library Website," *Tik Ilmeu J. Ilmu Perpust. dan Inf.*, vol. 7, no. 1, p. 105, Jun. 2023, doi: 10.29240/TIK.V7I1.6453.

[30] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *J. Big Data*, vol. 9, no. 1, pp. 1–36, Dec. 2022, doi: 10.1186/S40537-022-00592-5/TABLES/9.

[31] A. M. Deshmukh and R. Chalmeta, "User Experience and Usability of Voice User Interfaces: A Systematic Literature Review," *Inf. 2024, Vol. 15, Page 579*, vol. 15, no. 9, p. 579, Sep. 2024, doi: 10.3390/INFO15090579.

[32] U. Oh, H. Joh, and Y. Lee, "Image Accessibility for Screen Reader Users: A Systematic Review and a Road Map," *Electron. 2021, Vol. 10, Page 953*, vol. 10, no. 8, p. 953, Apr. 2021, doi: 10.3390/ELECTRONICS10080953.

[33] A. Oldeweme, J. Märtins, D. Westmattelmann, and G. Schewe, "The role of transparency, trust, and social influence on uncertainty reduction in times of pandemics: Empirical study on the adoption of COVID-19 tracing apps," *J. Med. Internet Res.*, vol. 23, no. 2, p. e25893, Feb. 2021, doi: 10.2196/25893.

[34] F. Tang and B. M. Østvold, "Transparency in App Analytics: Analyzing the Collection of User Interaction Data," *2023 20th Annu. Int. Conf. Privacy, Secur. Trust. PST 2023*, Jun. 2023, doi: 10.1109/PST58708.2023.10320181.

[35] A. Aldrees and D. Gračanin, "UX in E-government Services for Citizens: A Systematic Literature Review," *J. User Exp.*, vol. 18, no. 3, pp. 133–169, May 2023, Accessed: Sep. 27, 2025. [Online]. Available: https://uxpajournal.org/ux-in-e-government-services-for-citizens-a-systematic-literature-review/?

[36] J. Hochstetter, F. Vásquez, M. Diéguez, A. Bustamante, and J. Arango-López, "Transparency and E-Government in Electronic Public Procurement as Sustainable Development," *Sustain. 2023, Vol. 15, Page 4672*, vol. 15, no. 5, p. 4672, Mar. 2023, doi: 10.3390/SU15054672.

[37] B. Lu and J. Wang, "How can residents be motivated to participate in waste recycling? An analysis based on two survey experiments in China," *Waste Manag.*, vol. 143, pp. 206–214, Apr. 2022, doi: 10.1016/J.WASMAN.2022.02.034.